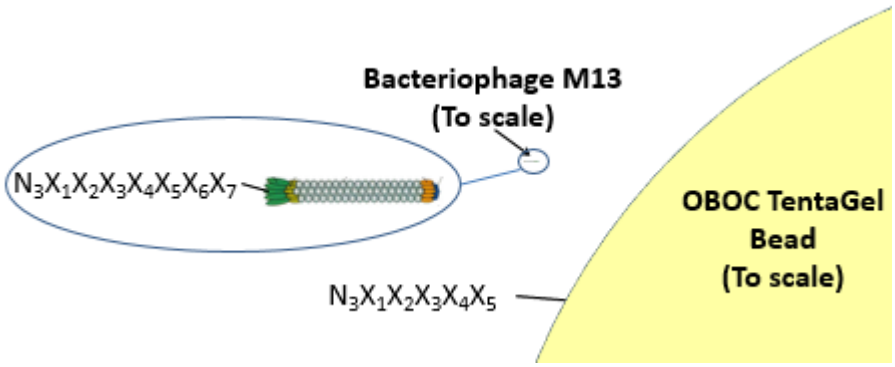
Chapter 5

Extending OBOC *in situ* Click Chemistry into a Phage Display System



5.1 Introduction

5.1.1 Azide-Containing Phage Display Libraries

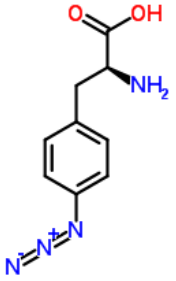


Figure 5-1: Azido - phenyl alanine *amino acid to be incorporated in phage library.*

Recent advances in chemical biology have made it possible to incorporate unnatural amino acids into recombinant proteins[1]. Schultz *et al.* has shown that through the use of amber suppression, azide-containing amino acids can be incorporated at specific locations into the pIII coat protein on an M13 phage[2]. The goal of this project is to synthesize a randomized 7-amino acid phage display peptide library containing an azide amino acid at a fixed position coded by an amber stop codon (TAG) to translate the click product screening technology into a biological library screening, and to demonstrate the use of resulting azide-phage library in *in situ* click screening.

The 7-randomized amino acid library will dramatically increase the number of peptide sequences that can be panned in each screen. For comparison, a complete OBOC library of 5 amino acids is sampled in 660 mg of library beads. To sample each sequence one time in a library the size of the 7-mer, it would require 447g of beads –-a practically impossible task using our current methods! In the phage display screen, however, it is trivial to completely oversample this 7-mer library 100 times in each screen–a huge advantage of the biological display technique. The strategies for discovering PCC agents with these libraries should be nearly identical to those previously developed for the OBOC click libraries, except that a much greater sequence space can be sampled with each screen due to the physical size of the library, and screening and sequencing steps can be performed significantly faster.

5.1.2 Mirror-Image Phage Display

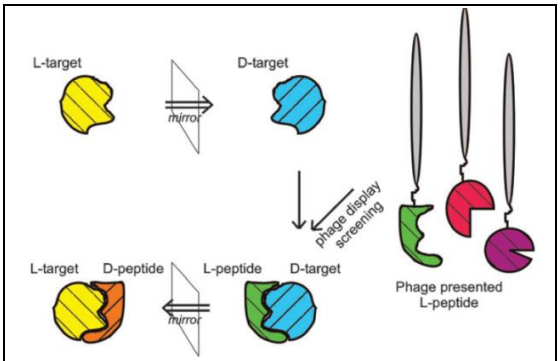


Figure 5-2: Mirror-Image Phage Display Technique[3].

One of the disadvantages common to phage display is that the library will contain exclusively L-amino acids, and will therefore produce a peptide sequence that is susceptible to protease cleavage. In order to avoid this potential problem, a technique called "mirror-image phage display" (**Figure 5-2**) has been developed. In this technique, an L-amino acid phage library is screened against a target synthesized from D-amino acids in order to find a D-peptide that binds to the L-target[4], where the D-target forms an exact mirror image of the L-target. Therefore, screening against a mirror image of the target and reversing the stereochemistry of the hit peptide binder produces a D-ligand that binds to the original L-target. With the epitope-targeting strategy currently used for PCC agent discovery in the lab, where a chemically-synthesized portion of the protein is used as a target for screening, it is trivial to prepare a D-amino acid epitope for use in a mirror image phage display screen. Using this technique, an azide-containing phage library can be screened in order to develop D-amino acid PCC agents, effectively converting the click screening so crucial to our success in PCC ligand discovery from an expensive and time-consuming OBOC library method to a simpler, quicker biological screening process.

5.1.3 G6PD Capture for Malaria Eradication

The Bill and Melinda Gates foundation is calling for the eradication of malaria, but this is a serious challenge that can only be addressed through the elimination of asymptomatic and chronic infections[5]. There exists currently a family of drugs, the 8–aminoquinolines such as primaquine, which can completely clear a person of infection (termed a "radical cure" regimen) and can thereby reduce the transmission of the disease[6]. Unfortunately, people with a glucose–

6–phosphate dehydrogenase (G6PD) deficiency, the most common enzymatic deficiency in the world, risk severe and life-threatening reactions to the standard treatment with this medication, which in turn requires a significantly different dosing method to effectively cure patients with a G6PD deficiency[7]. Therefore, in order to successfully employ the primaquine "radical cure" strategy for malaria eradication, effective methods to rapidly determine a patient's G6PD activity need to be developed immediately so that it is possible to administer the appropriate dose of the medication as quickly as possible[6].

The G6PD is found in red blood cells, and assists in the formation of NADPH from $NADP^+$, conferring protection from oxidative stress. Deficiencies normally arise from mutations in the G6PD gene that produce proteins with less than optimal function, reducing the overall enzymatic activity[6]. There are about 140 known mutations of this protein, most of which are single base changes, which can adversely affect the G6PD activity[7]. Current gold standard assays measure a deficiency by performing enzymatic tests on blood samples, measuring the amount of converted NADPH per unit of blood. Unfortunately, patients who have recently undergone a hemolytic event, causing a mass death of old red blood cells, show a false normal test. This is due to the higher than normal prevalence of young red blood cells, which express a higher G6PD copy number than mature red blood cells, compensating for the reduced activity of the enzyme caused by the mutation[6].

The goal of this project is to develop a capture agent that universally detects G6PD in all of the possible mutant forms in order to capture the protein to a chip where the concentration and the activity of the G6PD in the blood can be measured simultaneously. Such a diagnostic test can normalize the G6PD activity to protein copy number as opposed to activity per unit of blood, resulting in much fewer false positives and faster treatment of malaria. This combination test

should help reduce the number of false normal blood test results, allowing for significant progress toward the eradication of malaria.

5.2 Materials and Methods

5.2.1 Preparation of Plasmid for Incorporation of Azidophenylalanine and Amp Resistant Gene

The plasmid pAC-DHPheRS-6TRN (**Figure 5-3**) containing the coding sequences for *Methanococcus jannaschii* amber suppressor tRNATyr (MjtRNA) and the mutant *M.jannaschii* tyrosyl-tRNA synthetase (MjTyrRS) was a gift from Dr. Zhiwen Zhang at Santa Clara University. This plasmid was originally designed for the site-specific incorporation of 3,4-dihydroxy-L-phenylalanine (DOPA); its mutant MjtRNA recognizes TAG as a codon, and the mutant MjTryRS recognizes DOPA as a substrate. It also harbors a tetracycline-resistance selection marker. Since the host *E. coli* strains used for the phage production and in subsequent studies carry a Tet^R marker in their F' episomes, the Tet^R gene in this plasmid was first replaced with a beta-lactamase (Amp^R) gene for ampicillin selection, to be compatible with the antibiotic resistance of the host. Nine point mutations based on the description by Tian *et al.*[2] were then introduced into the MjTyrRS synthetase so that the resultant synthetase recognizes the azidophenylalanine amino acid instead of the DOPA.

To perform the Tet^R -> Amp^R switch, the Amp^R gene was amplified from a pET-3a plasmid by polymerase chain reaction (PCR) using the primers shown in **Table 5-1.** The PCR product and the original Tet^R-containing suppression plasmid were digested with HindIII and EagI, and ligated to produce slow-growing colonies on an LB-agar plate supplemented with 100 µg/mL ampicillin. Final clones were confirmed by DNA sequencing (Laragen).

Table 5-1: Primers for AmpR Switch.

Primer	Sequence
Hind_bla_5	GCG AAG CTT TAA TGC GGT AGT TTA TCA CAG TTA AAT TGC TAA CGC AGT CAG GCA CCG TGT ATG AGT ATT CAA CAT TTC CGT GTC GCC C
Eag_bla_3	ATA CGG CCG TTA CCA ATG CTT AAT CAG TGA GGC ACC TAT CTC AGC G

Based on the previous work by Tian *et al*.[2], nine point mutations were introduced into the mutant MjTyrRS gene to change its substrate specificity from DOPA to azidophenylalanine: E25K, L32T, S67A, N70H, E107N, D158P, I159L, L162Q, and Q167A. All mutagenesis reactions were carried out using QuikChange Kit (Stratagene) according to the manufacturer's instructions. **Table 5-2** lists the primers used in the order of performed reaction. Each QuikChange reaction was independently verified for the correct sequence, and the final plasmid was completely sequenced to ensure correctness. The resulting final plasmid was named pAC-AzPherS-6TRN.

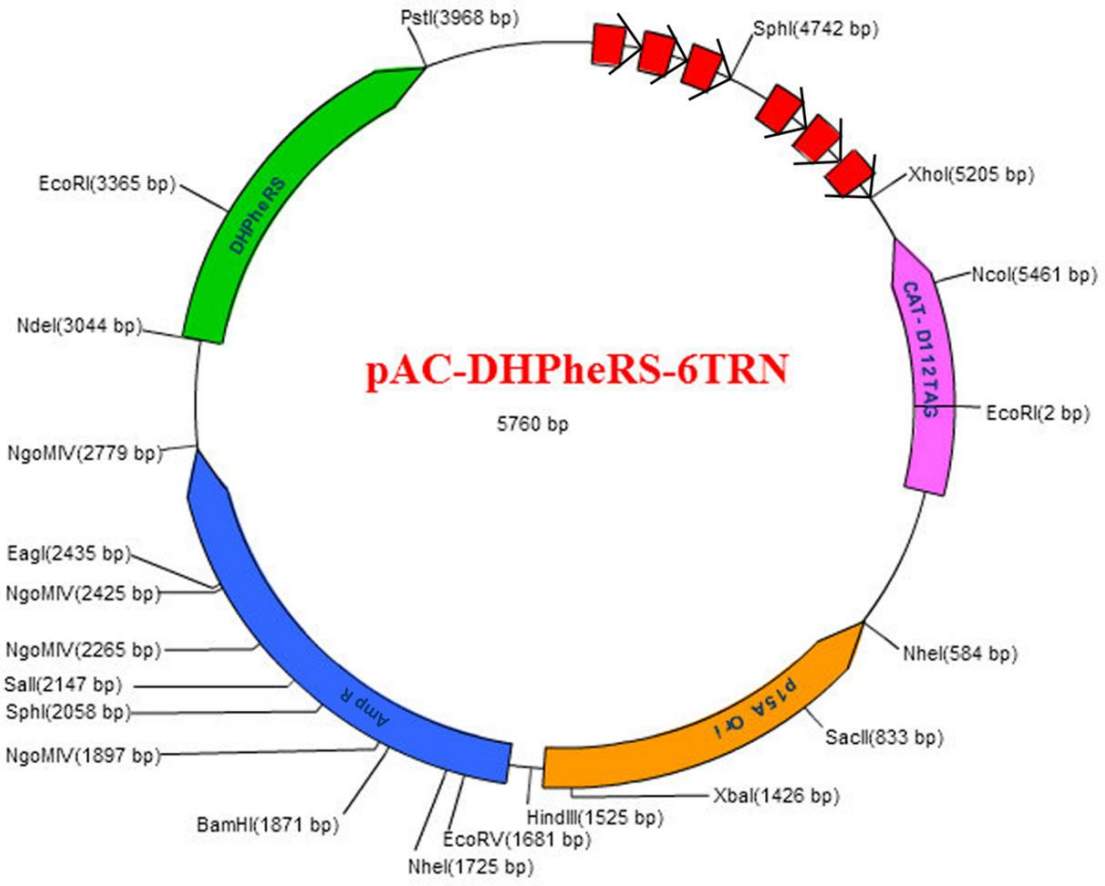


Figure 5-3: pAC-DHPheRS-6TRN Plasmid. *Plasmid for the incorporation of Azidophenylalanine. Red boxes are the six copies of the M. jannaschii tRNATyrCUA. Green indicates the synthetase. Blue is the Amp resistant gene, and the pink is the Chloramphenicol resistant marker containing an amber codon used to indicate successful incorporation.*

Table 5-2: Primers used for the DOPA -> Azidophenylalanine Synthetase QuikChange Reactions. *The number in the primer name indicates the order in which the QuikChanges were performed, and the mutation is included in the name of the primer.*

Primer Name/Mutation	**Sequence**
1_G73A_Upper	AGG AAG AGT TAA GAG AGG TTT TAA AAA AAG ATG AAA AGT CTG CTC T
1_G73A_Lower	AGA GCA GAC TTT TCA TCT TTT TTT AAA ACC TCT CTT AAC TCT TCC T
2_G319A_A321C_U	TTA AAG GCA AAA TAT GTT TAT GGA AGT AAC TTC CAG CTT GAT AAG GAT TAT ACA CTG
2_G319A_A321C_L	CAG TGT ATA ATC CTT ATC AAG CTG GAA GTT ACT TCC ATA AAC ATA TTT TGC CTT TAA
3_T199G_A208C_U	TGC TGG ATT TGA TAT AAT TAT ATT GTT GGC TGA TTT ACA TGC CTA TTT AAA CCA GAA AGG AG
3_T199G_A208C_L	CTC CTT TCT GGT TTA AAT AGG CAT GTA AAT CAG CCA ACA ATA TAA TTA TAT CAA ATC CAG CA
4_C94A_T95C_T96C_U	GAG GTT TTA GAA AAA GAT GAA AAG TCT GCT ACC ATA GGT TTT GAA CCA AGT GGT AAA ATA CAT
4_C94A_T95C_T96C_L	ATG TAT TTT ACC ACT TGG TTC AAA ACC TAT GGT AGC AGA CTT TTC ATC TTT TTC TAA AAC CTC
5_G472C_A473C_T474G_A475C_U	GGT TGC TGA AGT TAT CTA TCC AAT AAT GCA GGT TAA TCC GCT TCA TTA TTT AGG CGT CGA TGT
5_G472C_A473C_T474G_A475C_L	ACA TCG ACG CCT AAA TAA TGA AGC GGA TTA ACC TGC ATT ATT GGA TAG ATA ACT TCA GCA ACC
6_T484C_T485A_U	TAT CCA ATA ATG CAG GTT AAT GAT ATT CAT TAT CAA GGC GTC GAT GTT CAG G
6_T484C_T485A_L	CCT GAA CAT CGA CGC CTT GAT AAT GAA TAT CAT TAA CCT GCA TTA TTG GAT A
7_C499G_A500C_U	CAT TAT TTA GGC GTC GAT GTT GCG GTT GGA GGG ATG GAG C
7_C499G_A500C_L	GCT CCA TCC CTC CAA CCG CAA CAT CGA CGC CTA AAT AAT G

5.2.2 Test of Azidophenylalanine Incorporation into a Protein in E.coli

The ampicillin-resistant vector pAC-AzPheRS-6TRN developed above also carries a chloramphenicol acetyltransferase (CAT or Cm^R) gene where residue D112 is mutated to an amber stop codon. *E. coli* TOP10-F' cells were transformed with pAC-AzPheRS-6TRN plasmid and plated on an LB agar plate supplemented with ampicillin (100 μg/mL and tetracycline (12.5 μg/mL). Colonies were slow-growing, and were therefore incubated for 24 hours at 37C°. An LB broth culture of 5 mL supplemented with ampicillin (100 μg/mL) and tetracycline (12.5 μg/mL) was inoculated with a single colony and grown overnight at 37°C. The overnight culture (1 mL each) was diluted into two 7 mL cultures with appropriate antibiotics. An azidophenylalanine stock solution was prepared freshly by dissolving the amino acid in 50% DMSO and 50% acidic water (pH = 2.0 with HCl). A blank stock solution containing only the DMSO/acid without amino acid was also prepared. The amino acid (or blank) stock solution was added to the diluted culture to the final concentration of 2 mM (or 0 mM) azidophenylalanine. After incubating at 37°C for two hours, 34 μg/mL of chloramphenicol was added to each culture. Each culture was allowed to grow further at 37°C, and the level of growth was assessed by OD_{600} at 6 and 18 hours of incubation.

5.2.3 Test of Azidophenylalanine Incorporation into M13KE Phage

Four phage clones with variable display sequences of [Amber]-AHEATH, [Amber]-SHEATH, [Amber]-RHEATH, [Amber]–THEATH in M13KE phagemid were purchased from Antibody Design Labs. M13KE phagemid also contains the lacZα gene for blue/white plaque screening. A naturally-amber-suppressing strain of *E. coli* XL1 Blue was transformed with each individual clone, and plated on LA agar plate supplemented with tetracycline (12.5 μg/mL), IPTG (50 μg/mL) and X-gal (40 μg/mL) in top agar following the procedure described in Ph.D. Phage Display Libraries Manual (New England Biolabs, ref). An overnight starter culture of *E. coli* Top10-F' transformed with pAC-AzPheRS-6TRN was diluted 10-fold into 10 mL of LB supplemented with ampicillin (100 μg/mL)

and tetracycline (12.5 µg/mL) and grown at 37°C in the presence of 2mM azidophenylalanine for two hours. A blue plaque from each individual clone was added to the culture from the fresh plates, and cultures were further incubated for five hours. The cultures were centrifuged at 4500 rpm to remove the cells, and the phages were precipitated overnight at 4°C by collecting the top 8 mL of the supernatant and mixing it with 1.6 mL of 20% (w/v) PEG-8000 in 2.5M NaCl (PEG/NaCl). An alkyne-labeled TAMRA dye was clicked onto the azide moiety on the azidophenylalanine-containing phages using the Click-It Kit (Invitrogen) according to the manufacturer's instructions, and the resulting phages were resolved by SDS-PAGE. The gel was then imaged on a Typhoon imager using the preset settings for TAMRA dyes.

5.2.4 Synthesis of M13KE Azidophenylalanine-Terminated 7-mer Random Library

The peptide library was designed to have the following sequence when expressed on the surface of the pIII coat protein of the M13KE phage: A-[Amber]–$X_1X_2X_3X_4X_5X_6X_7$–G-L-V-P-R-G-S-pIII protein. The N-terminus of the display sequence has a fixed alanine amino acid to provide a non-charged spacer at the signaling peptide cleavage site to ensure proper expression of the azide-incorporated final pIII protein. The amber codon, the site of the azidophenylalanine incorporation, follows the N-terminal alanine, and precedes then the randomized 7-mer amino acid region. This library sequence is separated from the rest of the pIII protein by a thrombin cleavage site (LVPRGS) to allow the enzymatic cleavage of a "clicked" phage hit from a matrix-immobilized target.

The phage library was built by Antibody Design Labs using standard NNK codon, where N encodes an equimolar amount of cytosine, guanine, thymine or adenine bases, and K encodes only guanine or thymine. By designing the library in this fashion, some of the redundancy of the third codon as well as two of the three possible stop codons, TAA and TGA, are eliminated.

The final DNA library contained about 7.65 x 10^8 total transformants, and about 90% of that library contained inserts for a total sequence diversity of 6.9 x 10^8. It should be noted that there is a significant prevalence for guanine at the first position of a codon, as can be seen in **Figure 5-4**. This leads to a greater prevalence of the Glycine (G) amino acid in the overall distribution (**Figure 5-5**). Stop codons and Cysteine (C) amino acids are repressed since the production of functional phages is prohibited with these sequences, which is to our benefit. The statistical distribution of the library sequences is otherwise unremarkable.

Position	A	T	G	C	Total
First	84	60	128	57	329
Second	82	88	92	67	329
Third		176	153		329

Figure 5-4: Distribution of Random Nucleotides. *There is a significant prevalence for G in the first position, for unknown reasons. Recall that the third position is limited to T and G due to the NNK format of the library.*

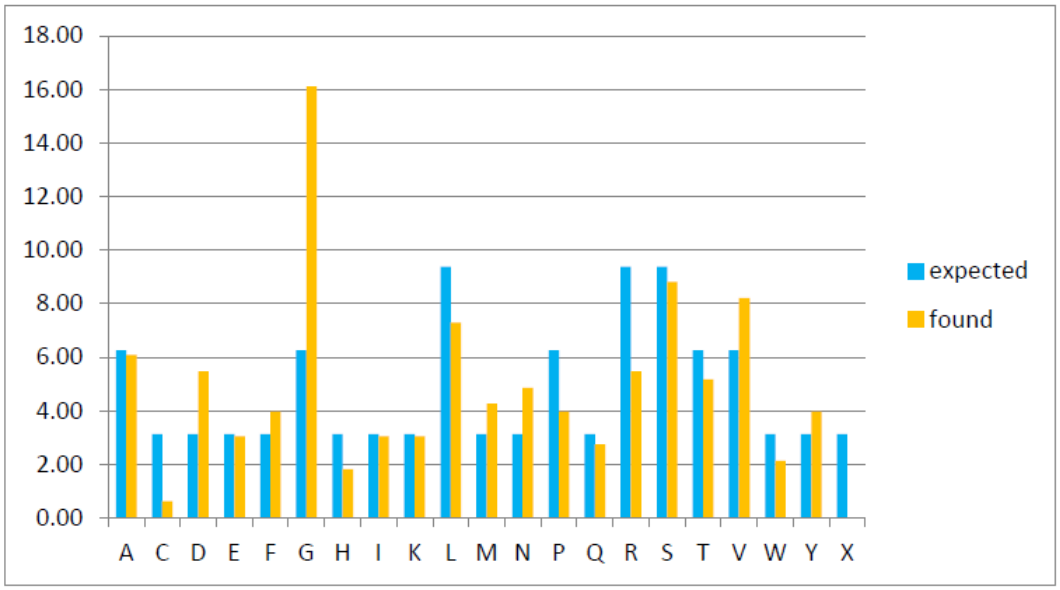


Figure 5-5: Distribution of Amino Acids. *The high prevalence of the G base pair in the first position lead to a much higher prevalence of the G amino acid than was expected. Otherwise, the distribution is as to be expected.*

5.2.5 Design and Synthesis of G6PD Target and Scrambled Target for Screening

The capture agent that will effectively detect G6PD in clinical settings must be able to bind to a site on the protein surface that has the least rate of mutation in order to ensure the most efficient capture of the protein with diverse sequence variance. As seen in **Figure 5-6**, exons 2 and 3 have the least number of total mutations identified in the polypeptide sequence. There are also no Class I mutations that correlate to the most severe deficiencies in enzymatic activity. Therefore, these two exons are the ideal location to target in order to capture the most diverse pool of mutant G6PD proteins. The region that is targeted by the capture agent should

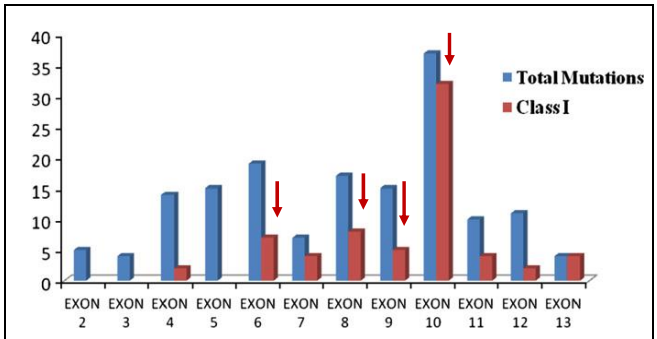


Figure 5-6: Location of G6PD Mutations in protein. *Exons 2 and 3 show the least number of mutations and no Class I mutations, so provide an ideal location to target.*

also be distant from the $NADP^+$ binding site in order not to interfere with the enzyme activity after capture. In the crystal structure (PDB ID: 1QKI) in **Figure 5-8**, the region encoded by exons 2 and 3 is unstructured in red in the bottom right of the protein. This region is far from the active site seen in complex with the $NADP^+$ at the top left of the protein, and the enzyme is expected to retain its original level of activity after being captured at this position.

The region including these exons, seen in **Figure 5-7**, has a few sites commonly mutated, which are shown in red. Therefore, the capture agent must be targeted to the amino acids shown in black in order to ensure it is not binding to a highly mutated section of the protein. The epitope that was selected for screening consists of amino acids 20 – 31 corresponding to the sequence LFQGDAFGQSDT, and was synthesized on TentaGel resin using D – amino acids in order to use the mirror image phage display technique to produce a D - amino acid peptide to bind to the L - G6PD protein. The synthetic target epitope was appended with propargylglycine (Pra) residue at the N-terminus, providing an alkyne click handle. A scrambled epitope also made of D – amino acids, [Pra]-GDAHSFQDTLQF, was synthesized on TentaGel as well, to be used in a preclear/antiscreen step for the library. The target and scrambled target on TentaGel resin will allow the phage library to click onto the permanently-immobilized target so that very harsh washing conditions can be used to remove all but the covalently bound phages. These peptide sequences were tested for their correctness and purity via Edman degradation sequencing.

Figure 5-7: Amino Acid Sequence of Exons 1 and 2 of G6PD. *The amino acids highlighted in red indicate frequently mutated positions, and the residues in black are generally conserved.*

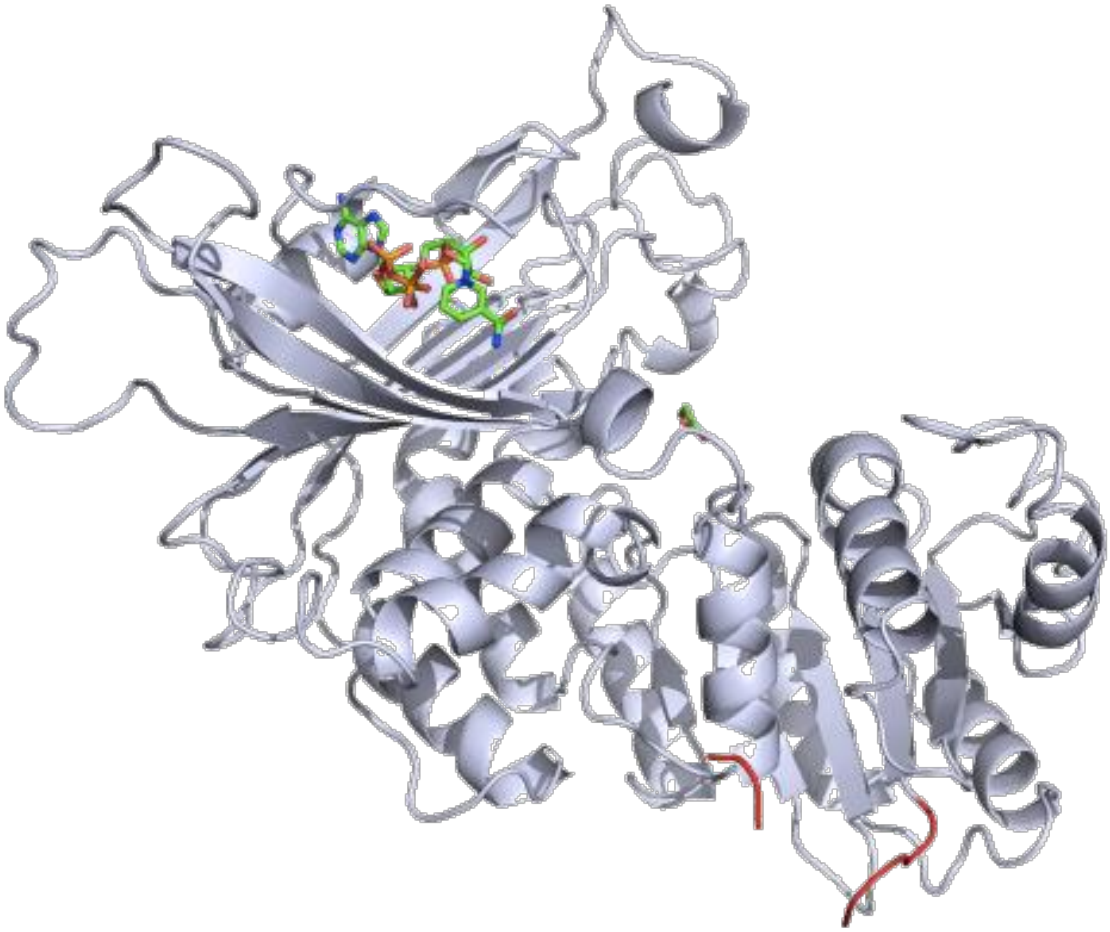


Figure 5-8: Crystal Structure of G6PD (1QKI). *This mutant G6PD in complex with NADP+ indicates that the unstructured regions encoded by exons 2 and 3 (bottom right in red, not entirely present in structure) are far from the active site of the protein. Capturing the protein at this location should not affect enzyme activity.*

5.2.6 Optimized Phage Library Target Screening Conditions

General protocols for phage library screening procedures including phage titering, scaling up, and sequencing as well as preparation of materials such as buffers, PEG/NaCl solution and top agar adopted from the NEB Ph.D. Phage Display Library System manual for M13KE, unless otherwise noted.

The library from Antibody Design Labs was produced in a naturally-suppressing strain of *E. coli* TG1 that inserts Glu, or E, at the site of amber codon. This library cannot be used for click screening, as there is no azide amino acid, but can be used for "target" screening to find peptide

sequences that bind to the G6PD epitope. The target screening must be completed first, also because approximately 35% of this phage pool contains no insert, or "naked phages." Naked phages have a significant selection advantage over those that would require the amber suppression. The target screen is designed, therefore, to enrich for phage containing the library insert, and remove the naked phage from the phage pool before incorporation of azidophenylalanine is performed.

The TentaGel beads containing the target and scrambled epitopes are stored in the form of 50% (v/v) slurry in 1:1 ethanol:water solution. One hundred µL each slurry of the target and scrambled epitopes were washed with 1 mL Tris-buffered saline (TBS, 25mM Tris-HCl, 150mM NaCl, 10mM $MgCl_2$) with 0.1 % Tween-20 (TBST), then blocked with 1mL of 5% (w/v) bovine serum albumin (BSA) in TBS for one hour at room temperature. An antiscreen was performed by incubating the scrambled peptide resin with a phage screening solution containing 2 µL sterile filtered human serum, 2 µL of the original library (~5×10^{11} total phages), and 196 µL TBS for 30 minutes at room temperature. The beads are spun down and the supernatant, containing the phage that did not bind to the scrambled peptide, is added to the target epitope resin and incubated for one hour at room temperature. Extensive washes are performed to ensure that the least number of non-specific and naked phages remain: five times with TBST, 30 minutes shaking in TBST, five times with TBST, five times with high salt TBS (2.5 M NaCl), 30 minutes shaking in high salt TBS, five times with high salt TBS, five times with TBST, five times with TBS, 10 minute shaking with TBS, five times with TBS, and two times with $CaCl_2$ thrombin buffer (1 mM $CaCl_2$ in TBS). The last wash step is retained and titered to ensure no phage are eluting. The target beads are then incubated with 4 units of thrombin in 200 µL $CaCl_2$ buffer for 24 hours at room temperature to elute hit phages. This step should elute only the phages that have a thrombin cleavage tag. After incubation, the 200 µL of thrombin cleavage cocktail was removed from the

beads, and the beads were washed five times with 200 µL TBS to remove all of the cleaved phages. The washes were combined with the thrombin eluent. The target beads are then acid-bumped by adding 1mL of 0.2 M glycine-HCl + 1mg/mL BSA buffer (pH = 2.0) for 10 minutes in order to ensure that all thrombin cleaved peptides were removed previously, and that only naked phage remain. The 1mL of acid eluent was removed from the beads and neutralized with 150 µL of 1 M Tris (pH = 9.1). Both the thrombin cleavage and acid bump eluents are titered to determine the number of phages eluted in each step.

5.2.7 Testing Phage Plaques for Library Inserts

The titered phage from a library or final eluent of a screen can be rapidly assessed for the presence of library inserts by plaque PCR before sequencing each individual clone from titered plaques. The first two primers listed in **Table 5-3** were designed by Antibody Design Labs, to recognize the DNA sequence on the M13KE phagmid on the either side of the cloning site, allowing the screening for the presence of insert based on the increased size of the PCR product compared to the one from a naked phage. The last PCR primer is an alternate 3' primer designed to recognize part of the insert DNA sequence, and therefore, a positive PCR reaction only occurs when the template phage contains the insert, and not with a naked phage. At each stage of the screening process, plaques were picked, and plaque PCR was performed to determine the percent of naked phage.

Table 5-3: Primers for Colony PCR of Inserts. *The first two primers show a slightly heavier PCR band if the insert is present. The last primer is an alternate 3' primer designed to sit inside the insert and, therefore, the PCR will show nothing if the insert is not present.*

Primer Name	Sequence
M13gV_5	GTC AGG GCA AGC CTT ATT CAC TG
Psi_3	GCG TAA CGA TCT AAA GTT TTG TCG
Thrombin_3	CGA ACC ACG CGG AAC CAG AC

5.2.8 Incorporation of Azidophenylalanine into Phage Libraries

The TOP10-F' cells transformed with the suppressor plasmid pAC-AzES-6TRN were streaked from a glycerol stock on an LB agar plate supplemented with ampicillin (100 µg/mL) and tetracycline (12.5 µg/mL). A single colony was inoculated into LB broth with appropriate antibiotics, and grown overnight at 37°C. The starter culture was diluted 100-fold into 20mL of LB + 12.5 µg/mL Tet + 100 µg/mL Amp, and allowed to grow for two hours at 37°C to early log phase (OD_{600} = ~0.1). A stock solution of azidophenylalanine was freshly prepared by dissolving 8 mg of amino acid in 250 µL of DMSO and 250 µL of acidic water (pH = 2.0) The entire volume of the stock azidophenylalanine solution was added to the culture to achieve the final concentration of 2 mM. For a control culture with no amino acid, only the 250 µL DMSO and 250 µL pH = 2.0 water are added. The cultures were then allowed to grow for additional one hour before the phage stock was added. For the target screen hits, the entire 1.2 mL phage elution was added to the growing culture. For the previously amplified library, phage amounts of 100 times the library diversity were added. The phage culture was then allowed to grow for 5 hours, centrifuged at 4500 rpm to remove the cells, and the top 80% of the supernatant was collected and precipitated with 1/6th volume of PEG/NaCl overnight at 4°C. The mixture containing the precipitated phage was centrifuged at 14,000 rpm for 15 minutes. The pellet was dissolved in 200 µL TBS and spun down again at 14,000 rpm to remove cellular debris. The total number of recovered phages was estimated before titering by measuring the absorbance at 269 nm and 320 nm with a NanoDrop spectrophotometer and using the formula in **Figure 5-9** from the Antibody Design Laboratories website[8], where the vector size was set to 7200bp for M13KE.

$$\frac{(A_{269}-A_{320}) \cdot 6\times10^{16}}{\text{number of bases/virion}} = \text{virions/ml}$$

Figure 5-9: Formula for Estimation of Phage Concentration.

5.2.9 *Optimized Phage Library Click Screening Conditions*

The click phage screens were performed on the library containing the azidophenylalanine after the incorporation step. These screens were performed exactly as the target screen, except that the phage library was incubated with the target on resin overnight at room temperature to allow sufficient time for the click reaction to occur. A diagram of this screen is seen in **Figure 5-10**.

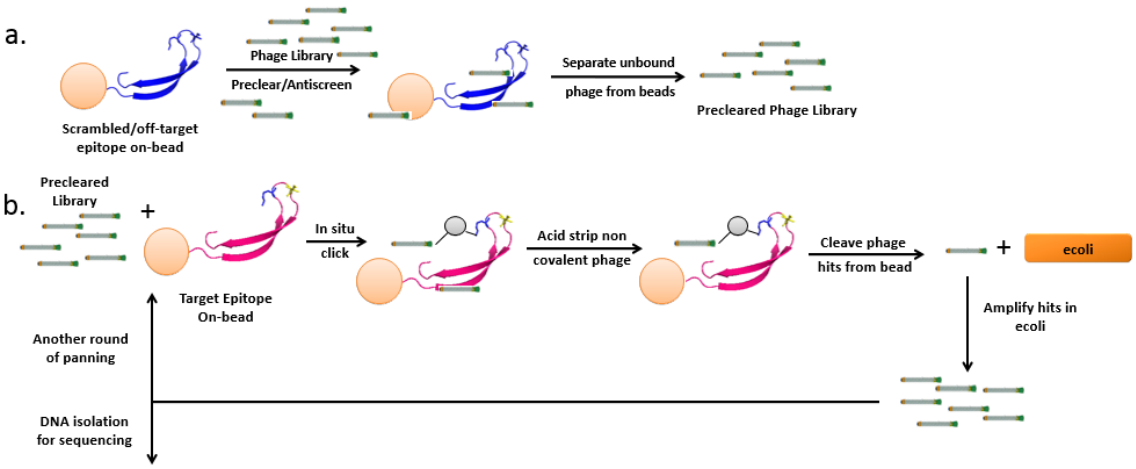


Figure 5-10: Phage Click Screening Strategy. *a.) The phage library is precleared for TentaGel and sticky binders by preincubating the library with a scrambled epitope on resin. b.) The phage are then incubated with the target overnight for a click to form, washed extensively, and eluted from the resin by thrombin cleavage for amplification in e.coli or titering and sequencing.*

5.3 Results and Discussion

5.3.1 *Test of Azidophenylalanine Incorporation into a Protein in E.coli*

The first test was to determine whether, and if so, in which condition the suppressor plasmid pAC-AzPheRS-6TRN created by the series of mutagenesis described in 5.2.1 is effective in incorporating azidophenylalanine into a recombinant protein in *E. coli* in response to an amber codon. Because the plasmid contains a chloramphenicol resistant marker (Cm^R) with an amber codon in its coding sequence (CAT-D112TAG), this antibiotic resistance is only conferred upon successful suppression of amber codon to synthesize the full-length protein. When grown in the presence of chloramphenicol in a liquid medium, only the culture containing the

azidophenylalanine amino acid was able to survive, as seen in **Figure 5-11**, with measured ODs of 0.639 with added azidophenylalanine as opposed to 0.108 with control. This result indicates that the mutated MjTyrRS for azidophenylalanine is functional and incorporating azidophenylalanine amino acid into the amber stop codon of the protein. It also indicates that the conditions used for this incorporation are sufficient to keep the background suppression without the unnatural amino acid to a negligible level.

Figure 5-11: Test of azide incorporation in e.coli.

5.3.2 Test of Azidophenylalanine Incorporation into M13KE Phage

After the successful incorporation of the azidophenylalanine amino acid into a protein in *E.coli*, the incorporation was tested using a small number of M13KE phage clones that contain representative display peptide sequences similar to the library that was designed for the subsequent *in situ* click screening. The azidophenylalanine incorporation was performed using four phage clones individually, and the presence of azide moiety in the collected phages was tested by clicking on an alkyne-containing TAMRA dye using a Click-It Kit (Invitrogen). This allowed for the fluorescence visualization of the dye-labeled pIII protein resolved by SDS-PAGE. As seen in **Figure 5-12**, there is one band across three of the four lanes corresponding to ~65kDa, or the running weight of the pIII phage coat protein, indicating that three of the four phage clones showed incorporation of the azidophenylalanine. The fourth phage could have been incorporating improperly or could have been lost in the multiple steps of the experiment, since the quantities of phage used were very small. Also, there are no other bands present in the gel which would indicate nonspecific azide incorporation in other proteins in the phage. This suggests that the incorporation of

azidophenylalanine is specific to the pIII protein where amber codon is introduced, and that peptides containing azidophenylalanine at a precise location can be displayed on the phage surface.

5.3.3 Phage Library Screening Conditions and Results

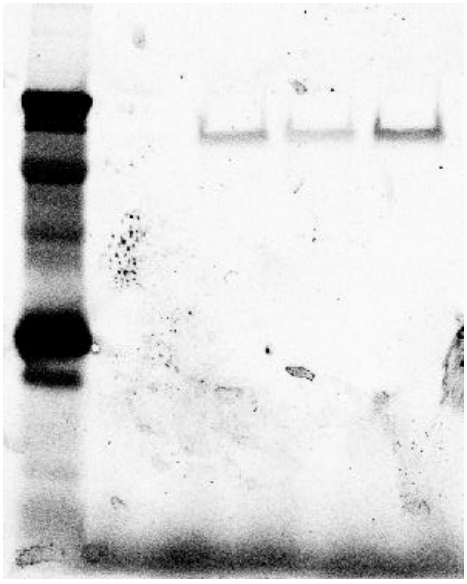


***Figure 5-12: Click-It Kit Visualization of Azide in pIII Coat Protein**. The band seen above corresponds to ~65kDa, the known running mass of the pIII protein.*

The original phage library produced by Antibody Design Labs contained about 35% naked phage determined by titering followed by plaque PCR, and cannot be directly carried onto the azidophenylalanine incorporation. The high prevalence of naked phage in the library complicates the subsequent azidophenylalanine incorporation, as the rapidly infecting naked phages that do not require suppression of the amber codon have a significant growth advantage over the properly-inserted phages, and can overwhelm the relatively slow synthesis of azidophenylalanine-containing phages by amber suppression. For this reason, the traditional phage display screening methods–"target" screens searching only for binding of the library component to the target of interest–were used to minimize the number of naked phages present in the incorporation step. The target screens eluted with thrombin cleavage were able to reduce the total amount of naked phage seen on titer plates to about 18%. Unfortunately, when this pool of eluted target hits was amplified using a naturally-suppressing *E. coli* strain XL1-Blue, this number rose to 85%-too high to use for any incorporation. An enrichment strategy that can reduce the number of naked phage to a manageable level has yet to be achieved.

5.3.4 Focused Library Screening

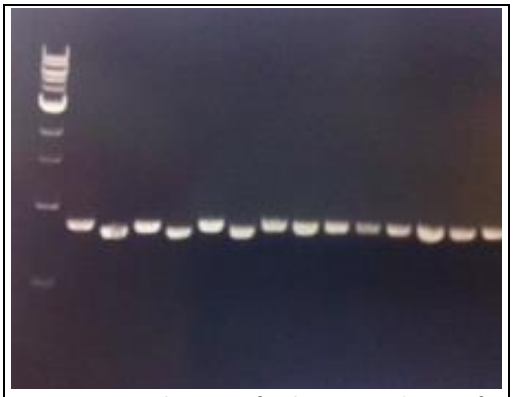


Figure 5-13: Gel Image of Colony PCR*. This test for library insert shows only 3 naked phage, lanes 2, 4, and 6. Out of 16 plaques total (not all shown on this gel), 13 contained an insert, and were used for the "focused" library.*

Because the amount of naked phages cannot be reduced to a level that would allow for successful azidophenylalanine incorporation, a "focused" library was designed from the insert-containing hit phages isolated from a target screen. **Figure 5-13** shows a result from the plaque PCR that was completed after a target screen to search for phages containing inserts. In this gel, each lane represents a single hit clone, and lanes 2, 4, and 6 show naked phages, as indicated by the PCR product of lower molecular weight. The phage plaques that were tested in this PCR were also individually amplified in 1mL cultures in order to create a small pool of each sequence. The 13 insert-containing plaques were then sequenced to ensure that each one contained an insert. As can be seen in **Table 5-4**, only 11 of the 13 hits contained one clean sequence. The rest appeared to have multiple sequences and may have had some naked phage contamination, and were thus eliminated from our pool of potential hits. Therefore, out of this target screen, 11 of the amplified stocks that contained inserts (**Table 5-4**) were pooled to create a "focused" library of phage that are known to bind to the target. This focused library has the added advantage of containing no naked phage, so it is the ideal library to be used to test the conditions for the azidophenylalanine incorporation. It will also be used for the subsequent focused library click screen with the target to test the *in situ* click on a phage library.

Table 5-4: Hits from 13 Insert-Containing Phages, Figure 5-13

#1	Amber	D	A	L	L	P	T	V
#3	Amber	N	S	T	Y	A	N	S
#4	Amber	I	S	A	Y	L	I	Q
#5	Uncallable, multiple sequence							
#6	Amber	A	F	S	A	L	D	L
#8	Amber	M	L	V	P	L	K	P
#10	Amber	M	D	T	W	L	M	T
#11	Amber	T	L	M	G	Q	W	W
#12	Amber	S	Y	T	T	M	E	V
#13	Amber	G	V	G	G	P	G	P
#14	Uncallable, multiple sequence							
#15	Amber	E	W	W	P	G	V	W
#16	Amber	V	L	H	G	G	R	A

5.4 Conclusions

The difficulties associated with the original library itself have delayed the testing of the described technology, though a roundabout solution to that problem has been discovered. The focused library is currently undergoing testing as the proof-of-concept for the click phage screening.

Future libraries for azidophenylalanine incorporation should be built using a phage vector that is out of frame, so that the insert is required in order to correctly synthesize the pIII coat protein. This will entirely eliminate the problem with the naked phage, and the complicated screening strategies will no longer be as important.

5.5 Acknowledgements

This ongoing work is being done in conjunction with Dr. Aiko Umeda. JingXin Liang performed the background research on the G6PD deficiency in malaria, and assisted in the selection of the screening target.

5.6 References

1. Wang, L.; Xie, J.; Schultz, P. G., EXPANDING THE GENETIC CODE. *Annual Review of Biophysics and Biomolecular Structure* **2006,** *35* (1), 225-249.
2. Tian, F.; Tsao, M. L.; Schultz, P. G., A phage display system with unnatural amino acids. *J Am Chem Soc* **2004,** *126* (49), 15962-3.
3. Funke, S. A.; Willbold, D., Mirror image phage display-a method to generate d-peptide ligands for use in diagnostic or therapeutical applications. *Molecular BioSystems* **2009,** *5* (8), 783-786.
4. Schumacher, T. N. M.; Mayr, L. M.; Minor, D. L.; Milhollen, M. A.; Burgess, M. W.; Kim, P. S., Identification of d-Peptide Ligands Through Mirror-Image Phage Display. *Science* **1996,** *271* (5257), 1854-1857.
5. Liu, J.; Modrek, S.; Gosling, R. D.; Feachem, R. G. A., Malaria eradication: is it possible? Is it worth it? Should we do it? *The Lancet Global Health* **2013,** *1* (1), e2-e3.
6. Domingo, G. J.; Satyagraha, A. W.; Anvikar, A.; Baird, K.; Bancone, G.; Bansil, P.; Carter, N.; Cheng, Q.; Culpepper, J.; Eziefula, C.; Fukuda, M.; Green, J.; Hwang, J.; Lacerda, M.; McGray, S.; Menard, D.; Nosten, F.; Nuchprayoon, I.; Oo, N. N.; Bualombai, P.; Pumpradit, W.; Qian, K.; Recht, J.; Roca, A.; Satimai, W.; Sovannaroth, S.; Vestergaard, L. S.; Von Seidlein, L., G6PD testing in support of treatment and elimination of malaria: recommendations for evaluation of G6PD tests. *Malaria journal* **2013,** *12*, 391.
7. Cappellini, M. D.; Fiorelli, G., Glucose-6-phosphate dehydrogenase deficiency. *Lancet* **2008,** *371* (9606), 64-74.
8. Phage Concentration Calculator. http://www.abdesignlabs.com/technical-resources/phage-calculator/ (accessed May 2).